DOCUMENT RESUME

ED 179 250 IR 008 177

| | |
|---|---|
| AUTHOR | Daniel, Evelyn H.; Ely, Donald P. |
| TITLE | Assessing the Competencies of Media Professionals: A Model for Determining Costs and Effectiveness. |
| INSTITUTION | Syracuse Univ., N.Y. ERIC Clearinghouse on Information Resources. |
| SPONS AGENCY | National Inst. of Education (DHEW), Washington, D.C. |
| PUB DATE | 79 |
| CONTRACT | 400-77-0015 |
| NOTE | 68p.; For a related document, see ED 149 740 |
| AVAILABLE FROM | Syracuse University Printing Services, 125 College Place, Syracuse, NY 13210 (IR-39; $4.75 prepaid) |
| EDRS PRICE | MF01/PC03 Plus Postage. |
| DESCRIPTORS | *Competence; *Cost Effectiveness; Higher Education; *Media Specialists; *Performance Based Education; *Student Evaluation |
| IDENTIFIERS | Information Analysis Products |

ABSTRACT

This is a summary of the procedures and outcomes of a study undertaken to develop alternative strategies for evaluating the competencies of library media personnel in competency-based programs. The study was prompted by a realization that, although competency-based programs for school library media specialists are increasingly being adopted by state agencies as a basis for certification and by educational institutions as a procedure for educating media professionals, the effort in developing these programs has been heavily weighted toward the generation of competency statements with little attention being paid to the perhaps more difficult problem of competency assessment. The methodology for this project involved working back and forth between literature review and theoretical model development, and the collection of empirical data testing various aspects of the model with formal and informal interviews of various experts to critique the developing model. A cost-effectiveness approach was selected as the most appropriate method, with the ultimate goal of providing educational decision-makers with an array of assessment strategies for each individual competency, together with the costs and effectiveness measures associated with each strategy. (Author/RAO)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED179250

# ASSESSING THE COMPETENCIES OF MEDIA PROFESSIONALS: A MODEL FOR DETERMINING COSTS AND EFFECTIVENESS

by

Evelyn H. Daniel

and

Donald P. Ely

ERIC Clearinghouse on Information Resources
Syracuse University

1979

IR008177

Evelyn H. Daniel is an Associate Professor in the School of Information Studies with a courtesy appointment in the Program in Instructional Design, Development, and Evaluation in the School of Education, Syracuse University.

Donald P. Ely is a Professor in the Program in Instructional Design, Development, and Evaluation in the School of Education, Syracuse University, and Director of the ERIC Clearinghouse on Information Resources.

This publication was prepared with funding from the National Institute of Education, U.S. Department of Health, Education, and Welfare under contract no. NIE-400-77-0015. The opinions expressed in this report do not necessarily reflect the positions or policies of NIE or HEW.

TABLE OF CONTENTS


| | Page |
|---|---|
| Competency Assessment Strategies | 1 |
| Competency-Based Education for School Library Media Specialists | 4 |
| The State of the Art of Competency Assessment | 7 |
| Classification and Definitions of Assessment Strategies | 12 |
| Measures of Effectiveness | 16 |
| Primary Effectiveness Analysis | 17 |
| Secondary Effectiveness Analysis | 21 |
| Measures of Efficiency--Cost Analysis | 23 |
| Primary Cost Analysis | 23 |
| Assessment Function | 26 |
| Resources | 28 |
| Types of Costs | 31 |
| Basis of Cost Comparison | 32 |
| Time and Usage Estimation | 34 |
| Secondary Cost Analysis | 35 |
| The Cost-Effectiveness Model | 39 |
| The Matrices | 41 |
| Data Display for Decision-Making | 42 |
| Fixing Costs or Fixing Effectiveness | 43 |
| General Competency Assessment by Faculty | 45 |
| Distribution of Competencies within Courses | 45 |
| Distribution of Assessment Strategies within Courses | 48 |
| Cost and Effectiveness of Course-Related Assessment Activities | 50 |
| Elaboration of Cost-Effectiveness Model | 54 |
| References | 55 |

LIST OF ILLUSTRATIONS

Page

Fig. 1 - Assessment Process Model . . . . . 3

Fig. 2 - Data Collection Categories for Competency Assessment Techniques . . . . . 13

Fig. 3 - Effectiveness Analysis Matrix . . . . . 18

Fig. 4 - Life-Cycle Costs . . . . . 24

Fig. 5 - Primary Cost Analysis Matrix . . . . . 27

Fig. 6 - General Cost-Effectiveness Analysis Model for Competency-Based Education Assessment Programs . . . . . 40

Fig. 7 - Cost-Effectiveness Data Display for Decision-Making . . . . . 44

Fig. 8 - Course by Competency Matrix . . . . . 47

Fig. 9 - Frequency of Competency Assessment through Courses . . . . . 49

Fig. 10 - Cost Data Collected from Faculty . . . . . 52

Fig. 11 - Cost of Assessment-Related Activities: Exploratory Study . . . . . 53

## COMPETENCY ASSESSMENT STRATEGIES

This is a summary of the procedures and outcomes of a study undertaken by Drs. Daniel and Ely, entitled "Developing Alternative Strategies for Evaluating Competencies of Library/Media Personnel in Competency-Based Programs."

The study was prompted by a realization that although competency-based programs for school library media specialists are increasingly being adopted by state agencies as a basis for certification and by educational institutions as a procedure for educating media professionals, the effort in developing these programs has been heavily (indeed, almost exclusively) weighted towards the generation of competency statements with little attention being paid to the perhaps more difficult problem of competency assessment. The American Association for School Librarians (AASL) Certification Committee observes that

> ...the art of competency assessment is woefully behind... Our lack of experience in the task of performance assessment, the identification of data gathering techniques--including what data will be required and under what conditions, the validation of assessment criteria, the need for competent evaluators to implement the assessment process, the cost of candidate assessment in terms of time and dollars, and reaching consensus on the levels of certification all pose serious concerns for the profession. (Candidate Assessment Process for Professional Media Personnel, 1977)

To some extent, the general lack of attention to evaluation is an artifact of the sequential process of development. Competency-based education for school media specialists is still a new area. Competencies must first be identified and agreed upon before there can be any concern with whether or not a person has achieved a given competency

or set of competencies. The AASL Assessment Process Model makes the sequential process clear by beginning in matter-of-fact fashion with "Competence Specification." (See Figure 1 - Assessment Process Model)

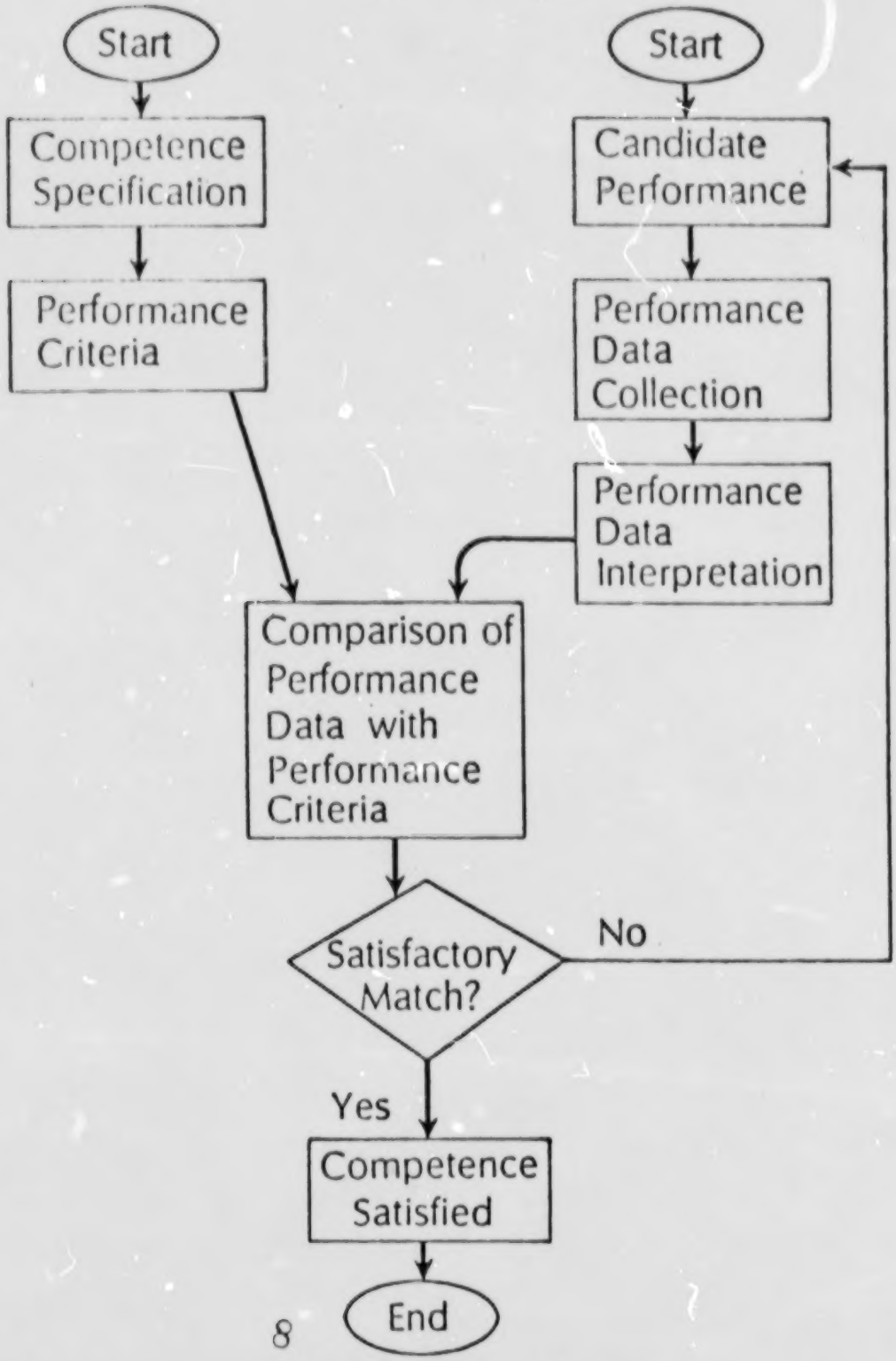
For the authors there was a natural and logical sequence of events leading to a concern for assessment. During 1977 a plan was created and carried out at Syracuse University for the development of an integrated competency-based program for the preparation of media professionals in response to a mandate from the New York State Education Department. A description of this process has been provided in the publication, A Process for Developing a Competency-Based Education Program for Media Professionals published by ERIC in 1978 (ED 149 740).[1] In this document, the techniques for identifying competencies and generating consensual agreement on their stated form and on their level of specificity are clearly and carefully delineated. The realization that the major work lay ahead in wrestling with assessment methodologies became clear as a result of that project.

A preliminary survey of the literature on competency-based education outside the media field showed that the lack of clear guidelines for developing assessment strategies was also true of all other areas where competency-based education was being implemented. This led to the development of a proposal

> to develop an assessment model that will provide a high degree of flexibility for students in the most parsimonious manner possible in order to best use the limited resources of the school.[2]

A series of specific objectives were put forward and a plan to accomplish their achievement was proposed. In July of 1978 this project was funded by the U.S. Office of Education through the Bureau of Libraries

# ASSESSMENT PROCESS MODEL



and Learning Resources, and the project was begun.

The methodology for this project involved working back and forth between literature review and theoretical model development and the collection of empirical data testing various aspects of the model with formal and informal interviews of various experts to critique the developing model. A cost-effectiveness approach was selected as the most appropriate method. The primary target audience for the study was the education community. The ultimate goal was to be able to provide educational decision-makers with an array of assessment strategies for each individual competency with costs (both monetary and other) and effectiveness measures associated with each strategy so that the educational decision-maker could make an informed choice. The intent was to demonstrate trade-offs clearly.

## Competency-Based Education for School Library Media Specialists

Competency-based education originated as an improvement of standard teacher training programs. It has spread to other professional education programs, such as nursing, social work, pharmacology, and library/media service. CBE as a curriculum innovation attempts to develop skills and attitudes which are essential for performing a set of specified tasks which have been identified as necessary for success in a given profession. Three basic assumptions underlie CBE.[3] First, for most jobs, it is possible to identify a finite group of observable behaviors (competencies) which are necessary and sufficient for competent professional performance. Second, appropriate performance measures for each competency or group of competencies can be designed and validated. Third, performance assessment should take place in settings

as like actual professional settings as possible.

In part the emergence and wide-spread acceptance of CBE derives from a perceived overbalancing in education towards the purely theoretical. Education has been noted for the continuing tension between theory and practice, between thinking and doing. A theoretical education deals in abstraction. The theoretician sees knowledge as an end, valuable for its own sake. A theoretical approach tends to be divergent, entertaining many possibilities and ideas. The nature of graduate education leads educators to value a theoretical approach more highly than a practical one which deals in concretes, sees knowledge as a means to an end, and tends to be convergent, focusing on improving proficiency in specified performances.

Obviously both approaches are important if one is educating a professional who must be able to perform adequately on the job, but who also must hold a larger view of the job and its environment and who must be prepared to adapt to unforeseen future changes. The CBE approach appears to be a response to this continuing theory-practice conflict that will pull professional education closer to the "real world" of practice. The danger, of course, and the concern that many educators have expressed, is the fear that graduate professional education may become overbalanced in the practical direction and that theory will be abandoned. Many of the developments in competency-based education in graduate schools can be traced to this concern.

There are basic differences between CBE as practiced in graduate education as opposed to undergraduate primary and secondary education. Most of these differences are a result of the need to blend the convergent method of CBE with the divergent environment of graduate edu-

cation. One difference is the more general level of objective specificity. A broad concept of competency is usually adopted for professional education rather than a narrow definition that spells out tasks in behavioral terms. The broad concept helps to focus on specific responsibilities which professionals will be expected to perform on the job. The narrow definition would not leave enough room for other learnings which are also important but difficult to specify.

Another difference is that the model of mastery learning is not assumed at the graduate level. Mastery learning asserts that the majority of students can learn the contents of a given curriculum if they are provided enough time to master the material at their own pace. Most professional schools adopting CBE do not accept the mastery model for the following reasons. First, there has already been a rigorous screening process designed to guarantee that those admitted to graduate programs are capable students and can accomplish the work presented in the curriculum. Second, because competition for admission can be high, those admitted should not be allowed an indefinite period to master material. Finally, professionals being produced by graduate schools are often placed in positions of responsibility immediately. Thus it seems neither reasonable from a school's perspective nor responsible from a broader societal view to allow students an extended period to reach mastery when repeated failure may indicate a lack of ability or of disposition for profeasional practice. As a result, CBE in graduate school is less an attempt to deal fairly with each student in allowing that student to learn at his/her own pace, than an attempt to certify that each graduate is capable of delivering professional service in the field upon successful completion of a course of study.

For similar reasons (especially for the social cost involved) graduate CBE usually does not include remediation opportunities nor place much emphasis on providing multiple learning approaches.

Finally, CBE at the graduate level is generally more idealistic and futuristic. The competencies of graduate training are designed to be both skills necessary at present in the field, documented through field practice, and competencies which ought to be used, but because of a lack of training or opportunity are not being demonstrated in current practice. Generally then, one can say that competencies in graduate school are required more to protect the clients of graduate professionals than to assure students of personal well-being or fulfillment.

These conclusions about the evolving nature of CBE in graduate programs have been reached in part by an extensive study of the writings on CBE and through interviews with those working in the area. The next section describes the current scene.

## The State of the Art of Competency Assessment

Major work on CBE assessment for school library media programs goes on apace. Eleven states now have competency-based certification programs and another nineteen are in various stages of development. Of these, Maryland (through the graduate library school), Utah (through the State Education Department in conjunction with several teacher training institutions), and New York (through the program at Syracuse University) have provided printed material of particular relevance. Both Maryland and Syracuse University work within existing curricula in large library schools which are only fractionally concerned with

the school media program. Thus both tend to take an incremental approach--identifying opportunities where students can attain competencies through general school courses, developing new courses, modifying older ones, and, in general, working with faculty styles of assessment. The State Education Department at Utah opted to support the development of a comprehensive pencil and paper testing program that established indicators for each competency and objective questions to assess student performance on the indicators.

Chisholm and Ely's book remains the major work in the field.[4] The American Association of School Librarians on the candidate assessment process for library media specialists[5] must also be considered a seminal work. The AASL group attempted to go beyond pencil and paper tests by providing examples of situations using jury panels of experts and actual performance activities.

A number of programs outside the media field, where advances in assessment techniques have been achieved, were identified. A brief mention of more notable programs follows. There were two approaches to assessing teacher competency, one in Oregon critiqued by Schalock,[6] the other in Georgia described by Okey et al.[7] Good reliability and validity studies accompanied the latter effort.[8] The undergraduate medical curriculum of the Southern Illinois University School of Medicine has been competency-based since its inception in 1973. A good summary of procedures with attention to the design of a competency maintenance system and comment on eight assessment procedures is provided by Williams.[9] The competency-based pharmacology program at the University of Minnesota employs an assessment center approach to competency measurement.[10] For school social workers, the University of Washington,

Seattle, features a consortium-directed competency-based program featuring a candidate review board that uses a peer review and advocacy process.[11] Ellis and Bryant describe some of the problems of evidence gathering and weighing inherent in this approach.[12] Alverno College, a small women's liberal arts college in Milwaukee, Wisconsin, awards its degrees on the basis of a CBE program using an assessment center concept.[13] The military have also made significant advances in assessment of competency through validation studies, behavior sampling, highly sophisticated simulation laboratories, and development of criterion-referenced/performance-based measures.[14] Especially interesting approaches were those reported at the 1979 AECT Conference by Worth Scanland of the Naval Education and Training Command,[15] Wayne Waag of the Air Force Human Resources Laboratory,[16] Robert Wilshire of the Army Signal Center,[17] and Marjorie Kupper from the U.S. Army Engineer School in Fort Belvoir, Virginia.[18]

General research on evaluation also provides some useful insights. DeProspo and Liesener describe current evaluation models in a concise but thorough review for media program evaluation.[19] Hall and Jones, in a general presentation of the theory and state-of-the-art of CBE raise and discuss twelve provocative questions affecting competency assessment.[20] Houston and Howsam review CBE with emphasis on teacher education,[21] while Hodgkinson et al. focus on assessment in the higher education area.[22] Harris and Kelly elaborate the higher education assessment work more fully.[23] Finally, two new books published within the last year demonstrate the substantial growth in the field. On Competence brings together review articles by experts on major aspects of CBE.[24] Of particular interest is the fine state-of-the-art review on

assessment by King.[25]

In the second recent book, Gilbert proposes a behavior engineering model.[26] He suggests that behavior has three aspects:

Information coming to the person telling him/her what to do
($S^D$ - the discriminative stimulus)

The person responding in some way
(R - response)

The action of responding and the stimulus being reinforced
($S^R$ - the reinforcing stimulus)

Performance also has two aspects: a person with a repertoire for behavior (P) and a supporting environment (E). Gilbert develops a simple 3x2 matrix as a way of examining competent performance with its antecedents more analytically. He suggests a new measure--the PIP. A PIP (Potential for Improving Performance) is the ratio of exemplary performance to typical performance. Through its use he identifies points where the opportunity exists to make the biggest positive change. Gilbert further specifies ways of measuring any performance through selected measures of quality (defined as accuracy, class, and/or novelty), quantity (rate, timeliness and/or volume), and cost (labor, material, and/or management). Gilbert's approach is an interesting extension of Skinnerian behavior modification and one that we would like to apply in future study of the field work component of CBE.

From the aforementioned and a number of other sources, it is possible to identify particular problems and approaches in the assessment of competency. In the first place, many professionals speak almost synonymously of evaluation and measurement. Actually, there is a growing consensus that evaluation is the broader term relating to programs and products more

than to individuals, whereas assessment refers to the specific evaluation of learning outcomes.[27] Rountree's delightfully literate discussion provides additional support for this distinction.[28]

Major problems continue with respect to precise definition of measures, determination of a "significant" sample of behaviors when collecting data on the measures, and, perhaps the biggest problem--validation of the measures.[29] Assessment centers are being established in many areas particularly for management development programs. A good overview on the potential for the assessment center approach is Applying the Assessment Center Method.[30]

A number of studies used some variation of a matrix as a systematic organizing tool because of the array of evaluation strategies, learning domain categories, presentation formats, etc. Hodgkinson says that CBE needs "instruments capable of much finer gradations, instruments capable of providing formative diagnostic advice, and instruments capable of telling us if the student is "real world" competent...."[31] They see the matrix as a device which facilitates triangulation (that is, different measures coming from different sources but all indicating the same result) greatly increasing the reliability of judgments based on the measure. Rogers uses a matrix to relate testing strategies in the field of library science and media.[32] Of course, one of the earliest to use a matrix strategy to present massive amounts of interrelated information in an orderly fashion was JIMS--the Jobs in Instructional Media Study.[33] The matrix is a device that we have adopted as useful for this study.

From this comprehensive and still on-going state-of-the-art review

as well as from our process of empirical data collection and sorting, it has been possible to develop a classification of assessment strategies which will be described in the next section.

## Classification and Definitions of Assessment Strategies

Early in the project, specific techniques for assessing competencies were collected in order to establish a data bank of assessment procedures for each of the seventy competencies currently in use for the Syracuse program, plus variations and additions from other CB programs throughout the country. To do this a standard collection instrument had to be devised. A list of the categories used to collect the information is shown in Figure 2.

The original expectation was that a comprehensive array of techniques would be relatively easily collectible. There was a plan to computerize the information for subsequent easy manipulation and retrieval and then to move quickly to testing the cost and effectiveness of alternative techniques. Although nearly four hundred specific techniques were collected, it would require many more techniques from a larger number of sources for the original approach to be a fruitful one.

This data bank was used to test and refine the category scheme for assessment strategies. From this approach eight strategies were identified which seemed to represent the most comprehensive, conventional, and feasible methodologies for assessment. The definitions of the eight assessment strategies are presented below.

COMPETENCY STATEMENT:

COMPETENCY NUMBER:

EXPLICATION:

COMPETENCY AREA:

LEVEL OF ASSESSMENT:

ASSESSMENT STRATEGY:

CONDITIONS:

CRITERIA:

INDICATORS:

ACTIVITIES:

SOURCE:

ACCESS TERMS:

Fig.2 - Data Collection Categories for Competency Assessment Techniques.

(1) Field Experience. Field experience refers to assessment which occurs in the real world context, that is, assessment which occurs in a real job situation or internship. Performance which occurs in this situation is subject to real world effects and tends to demonstrate competence at its most functional basis.

(2) Actual Demonstration. Actual demonstration is assessment which occurs in a field setting where many but not all of the real world effects are allowed to influence student performance. Since student access to field settings occurs through invitation or through agreement with cooperating institutions, some control exists to limit real world effects.

(3) Simulated Demonstration. A simulated demonstration refers to a microcosm type demonstration. In this case many of the variables affecting performance are controlled or manipulated. Effects due to change or real world contingencies play only a small part in the assessment situation.

(4) Portfolio. A portfolio is a collection of written materials, audio/visual, artistic, or craft products which are submitted as indication of field competence. A portfolio is generally composed of materials developed during the course of a program or some professional experience and is used to infer job related competence.

(5) Paper/Projects. Normally paper/projects refers to a product emerging from a major assignment. Examples include term papers, slide/tape productions, musical compositions, and so forth. Paper/projects may also refer to smaller assignments and projects. The idea here is that the work produced by the student is examined and assessed as an indicator of some knowledge, attitude, or skill obtained during instruction.

(6) Oral Presentation. Oral presentations are student-controlled summaries and discussions which reveal (1) the activities involved in, (2) the progress of, and (3) the findings related to student projects or papers. The oral presentations are more than impromptu conversations. They allow the audience to ask questions or make comments which may require the extension of ideas through in depth responses. Thus, such testing is more flexible than written presentations, though less detailed.

(7) Oral Test. An oral test is a strategy in which students respond to a set of questions specified by an examiner or group of examiners. This strategy allows

flexibility in pursuing answers more fully and clarifying responses, though this is generally at the expense of the organization and detail present in a written response.

(8) Written Test. Written tests are generally of two types: selected response (true-false, multiple choice), and constructed response (short answer, essays). There is virtually no interpersonal interaction in such tests, and competence is assessed on the one-way communication of the student with the written materials.

A separation into two subdivisions of (a) measures of competence and (b) indicators of competence, as illustrated below, seemed most appropriate

| | | | |
|---|---|---|---|
| (a) | (1) | Field Experience | Measures of Competencies |
| | (2) | Actual Demonstration | |
| | (3) | Simulated Demonstration | |
| (b) | (4) | Portfolio | Indicators of Competencies |
| | (5) | Paper/Projects | |
| | (6) | Oral Presentation | |
| | (7) | Oral Tests | |
| | (8) | Written Tests | |

Assessment strategies as measures are distinguished from strategies as indicators because measures assess actual performance in real or simulated job settings, while indicators assess behaviors thought to be more or less related to job performance. The assessment strategies of (1) field experience, (2) actual demonstration, and (3) simulated demonstration are said to measure competence because they attempt to record the existence of specified, necessary job skills and attitudes (i.e., the identified competencies). The assessment strategies of (4) portfolio, (5) paper/projects, (6) oral presentations, (7) oral tests, and (8) written tests are said to be indicators of competencies because they do not test actual performance either in simulated or actual job situations. Instead, from these strategies, one infers what performance might be like

in these situations.

This highly useful distinction between measures and indicators led to the problem of determining "good" techniques for assessing each competency.

## Measures of Effectiveness

Effective assessment outcomes are critical to a competency-based program. Determining the effectiveness of CBE assessment outcomes has been a problem in the CBE movement since its inception. Many of the problems may be due to the limitations of the art of measurement. These problems, however, may be more a matter of lack of experience with the technology of performance assessment than a true limitation.

Our approach to effectiveness analysis was first to ask the question: "To what degree do the assessment strategies accomplish what they set out to accomplish?" This is more or less equivalent to asking: Is the assessment strategy valid? Thus, in one sense, effectiveness analysis of assessment strategies will be identical to a validity analysis. This technique was selected as the most appropriate approach to performing a primary effectiveness analysis.

Validity analysis seems a necessary but not sufficient response to developing measures of effectiveness. A number of other questions are also important in considering the effectiveness of a given strategy. For example, how practical is it? Is there a reasonable expectation that the requisite mix of faculty experience and appropriate environment can be created? How complex are the directions for an assessment instrument? If the procedure is too complicated, will it affect the results? How

much effort is involved in constructing, administering, and scoring the assessment technique? While some people feel confident that primary effectiveness analysis can be made to yield good quantitative data, other kinds of data are also important in making judgments about how "good" a particular technique may be when used to assess a student's competency. Data on these qualitative questions are considered when performing a secondary effectiveness analysis.

The careful separation of quantifiable measures from qualitative measures clarifies the decision-making process without suggesting that one type of measure is any more important than another. Figure 3 demonstrates how the evidence on the comparative effectiveness for a set of strategies for each competency as well as for clusters of competencies can be arrayed in matrix form. What follows is a brief exploration of procedures for primary and secondary effectiveness analysis.

Primary Effectiveness Analysis. A basic assumption of CBE is that assessment strategies which examine performance provide the best measures of competence. The closer measures are to real world contexts, the more valid they tend to be. Strategies which are increasingly removed from real world contexts are potentially less valid as measures of competence. By this reasoning a simulated demonstration would be less valid than a field-based assessment (by how much is not known). Similarly a portfolio assessment would be less valid than a simulated demonstration. Validity is defined as the extent to which an assessment strategy does the job for which it is used.[34] When a strategy which has been designed to measure a competency actually does so, it is said to be valid. When an assessment strategy which has only the power to infer a

| | | CRITERIA | | | | |
|---|---|---|---|---|---|---|
| | | VALIDITY | RELIABILITY | OBJECTIVITY | PRACTICALITY | |
| | | | | | CONSTRUCTION | ADMINISTRATION | SCORING |
| STRATEGY | FIELD EXPERIENCE | | | | | |
| | DEMONSTRATION: ACTUAL | | | | | |
| | DEMONSTRATION: SIMULATED | | | | | |
| | PORTFOLIO | | | | | |
| | PAPER-PROJECT | | | | | |
| | ORAL PRESENTATION | | | | | |
| | ORAL EXAM | | | | | |
| | WRITTEN EXAM | | | | | |

Fig. 3 - Effectiveness Analysis Matrix.

competency is used to measure one, there is need to question its validity.

Effectiveness analysis of any assessment strategy is based upon one fundamental question: To what degree does the given assessment do what it is intended to do? Attempts to answer this question have been a primary focus for those in the field of educational and psychological tests and measurements. The research literature from this field can be of considerable assistance in resolving competency-based assessment problems. Holloway strongly states that the future success of CBE problems will depend upon what CBE educators know about the general field of tests and measurements.[35]

It is the concern of experts from the tests and measurement field to measure qualities, characteristics, skills, and/or competencies within an individual and to determine the effectiveness of their measures. The term "test" is usually considered to mean the presentation of a standard set of questions to be answered, the result of which is a score or numerical value which is said to represent the extent of a characteristic or a quality that a person possesses.

Test developers are not always satisfied that the tests they construct actually measure the characteristics they were designed to measure. This may be why they take great pains to determine the degree to which the tests actually measure what they are intended to measure. This is called test validation and requires an examination of the reliability of a test plus a closer examination of various types of test validity.

A test's reliability is that portion of the result (the score) that is due to systematic errors and persists, therefore, from sample to sample, or from one administration of the test to another.[36] Reliability

reflects the degree to which the test can deliver the same or similar scores with each testing, or deliver scores which reflect the true differences among students. Indications of reliability are presented in the form of coefficients of reliability. Reliability is considered a necessary condition for test validity but not a sufficient one.

The determination of whether or not a test measures what it is supposed to be measuring is examined in validity analysis. Validation is a procedure that provides "evidence that an inquiry is free from bias or otherwise conforms to its declared purpose."[37] In the absence of bias, we know that the results are not influenced by other factors and can be believed.

Test validity is not easily attained and at best is a matter of degree. Mehrens and Lehmann state:

> A test may have many validities, each dependent upon the specific purpose for which one uses the test. Eventually the validity of any test is dependent upon how it is used in the local situation.[38]

Test validation involves an examination of four types of evidence: (1) content, (2) mental constructs, (3) relevant criteria, and (4) appearance. Each of these forms of evidence is discussed briefly below in the manner typically employed by test and measurement experts.

(1) Content Validity. The test is examined to determine how the content of the test samples the domain about which inferences are made. An individual's test score can validly infer his/her knowledge of a subject only to the degree that the test has adequately sampled (i.e., provided representative examples of) the subject domain.

(2) Construct Validity. The test is examined in terms of the degree to which the test scores can be accounted for by certain constructs of a psychological theory.

It is interesting to reflect on how this notion might be replaced in CBE by competency validity which would examine the test in terms of the degree to which the scores can be accounted for by certain competencies.

(3) Criterion-Related Validity. The test is examined in terms of its relationship with external measures (which measure the same or similar things). For example, success in college (Grade Point Average) is indicated by certain other tests or measures obtained in high school (i.e., SAT, High School GPA). Such an examination speaks essentially to predictive validity. That is, the ability of a test score to predict other types of performance.

(4) Face Validity. Here the test is examined in terms of its ability to show how "on the face of it" that it measures what it says it measures.

The technology of test validation is a sophisticated one, based on a long history of theoretical and applied research. It is possible that this rich body of research can be mined to suggest ways in which test validation can be applied to forms of assessment other than the written test. This is the next major phase for research on competency assessment.

Secondary Effectiveness Analysis. A number of qualitative dimensions have been proposed in the literature that also purport to measure the effectiveness of assessment techniques. Examples of these dimensions include the determination of adequacy and appropriateness of assessment outcomes.[39] Doughty et al. suggest communicability, feasibility, and utility.[40] Knapp and Sharon add equitability.[41] Kelly proffers ethicality.[42] Meherens and Lehmann propose objectivity and practicality.[43] Cost, of course, is a major criterion, so much so that it has been separated from the other dimensions and treated as an independent category to be discussed in the next section.

Each of these characteristics pertains to some dimension of an assessment strategy which may influence the choice to keep or remove the strategy from the total assessment program. For example, equitability addresses the fairness of the assessment strategy. Scriven advocates an "inspecting process" which searches for injustices within the assessment strategy.[44]

> Whatever the outcomes, there are certain types of procedures (sic) that are inappropriate for moral reasons, and inspection of process must be made to see whether excess cruelty, inequitable methods of grading, etc., are important.[45]

Clearly where there is evidence that the assessment technique is really not fair to the student, educators are obliged to discard it, or to alter it until it is equitable.

Two problems are associated with secondary effectiveness analysis. One is the problem of what to consider when conducting secondary analysis. Since there are many characteristics suggested and since some even appear to conflict, selecting the appropriate ones can be a problem. However, this perhaps should be treated as a local problem to be decided according to the circumstances and conditions under which the analysis is taking place.

The second problem relates to definition. Most of the terms suggested above are too vague to apply reasonably. Even though the ultimate meaning of any of the terms will be determined by the location and conditions of analysis, a more precise and common set of definitions would be an important contribution. This also is part of the next phase of research.

Secondary effectiveness, however poorly defined at present, is

important because it imposes a realistic perspective on effectiveness-based decisions. There are always more things likely to effect the selection or rejection of an assessment's strategy than validity alone. Secondary effectiveness analysis assures a more comprehensive analysis.

A distinction is often drawn between measures of effectiveness and measures of efficiency. Efficiency questions focus on the consumption of time, money, and effort. Those concerns are addressed in the next section on cost analysis.

## Measures of Efficiency--Cost Analysis

In general, cost analysis attempts to discover the potential and/or actual costs of a program. As the term is used here, cost means the expenditures of human and monetary resources for the development and operation of programs. As it is currently conducted, cost analysis concentrates largely on the monetary resource expenditures. However, there is increasing attention given to the expenditure of human resources as well. This distinction between human and monetary costs provides the basis for a two-part analysis of costs in much the same way as that drawn above between primary and secondary effectiveness analysis. Monetary costs are quantifiable and will be used for the primary cost analysis. Human costs may be more qualitative in nature and will be the focus for secondary cost analysis.

Primary Cost Analysis. Cost, as defined by economists, is "the best alternative forgone."[46] However, cost will be used here in a more general sense as "a common and universal measure of the nature and quantities of resources used in a specific manner to achieve stated objectives."[47]

A universal standard is necessary for the comparison of alternatives (programs, assessment strategies, etc.) in order to choose among them. In cost analysis, the universal standard is usually dollars, and although there are some difficulties that will be discussed below, it seems the best reflection of cost available.

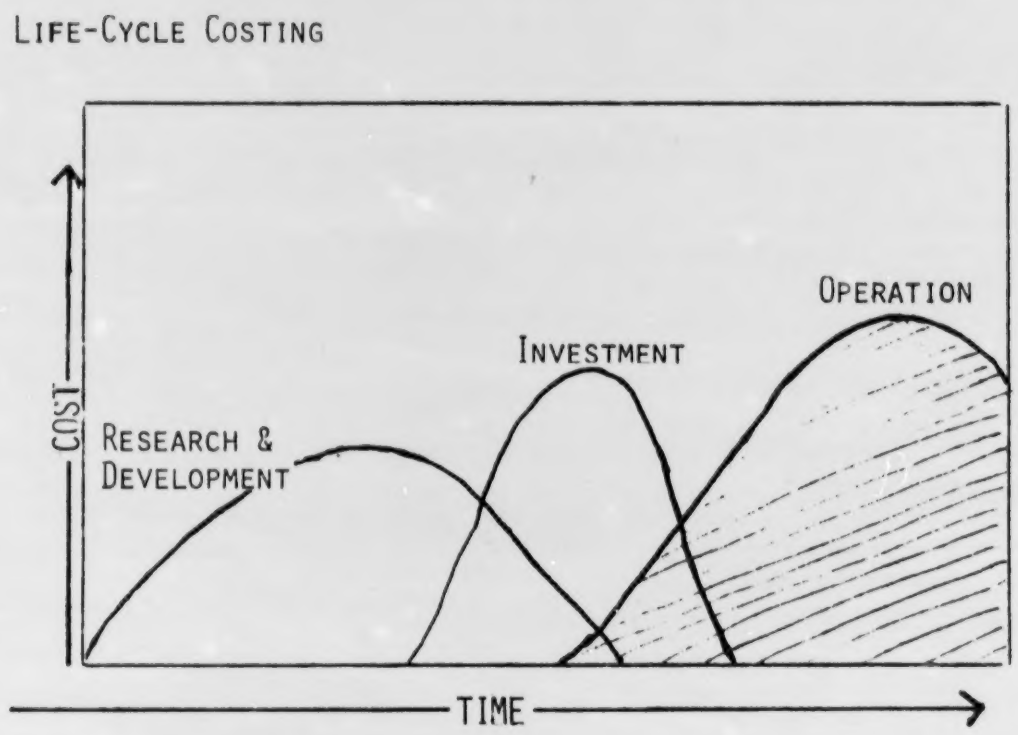



Fig. 4. Life-Cycle Costs

We have adopted a function by resource approach to costing following Doughty and Beilby,[48] Carpenter,[49] Stace,[50] and others. Program activities (functions) are analyzed in terms of the expenses they generate. These expenses are then distributed among the various categories of resources. The cost of these resources (personnel, equipment, facilities, material) is then estimated using standard prices for comparable resources. Methods of costing not based on the resources necessary for each individual or for each specific program or activity can suffer from one of two possible errors in estimation. First, costs may be included which would not be charged to any single program, because, for instance, they are shared among several programs. These are referred to as joint costs and will be discussed more fully later. Including joint costs entirely in a single program or activity thus inflates the apparent costs. The second error is the obverse of the first. Costs which are obscured by more traditional accounting and budgeting procedures may be omitted thus depressing the actual cost of a program. "(Function)/Resource cost analysis is designed to provide cost estimates which permit equitable comparisons among projects (strategies), where cost or cost/effectiveness is the basis of comparison."[51]

Although the cost model used by Doughty et al.[52] and others examined developmental as well as operational costs of programs, we will disregard the developmental costs and concentrate on operational costs. The shaded portion on Figure 4 indicates our focus. This is done in order to analyze costs along a third dimension--that of the eight assessment strategies. The cost analysis model, then, analogously to the effectiveness model, can be shown as a three dimensional matrix displaying operational costs generated by the interaction of functions with

resources with assessment strategies. Figure 5 graphically portrays these dimensions and their interrelationships.

The procedure begins with the identification of a set of functions that are mutually exclusive yet collectively include all the activities that take place in the assessment process. A parallel process involves the determination of all the resources used. Both function and resources categories must be carefully and explicitly defined so that an observer or a participant can tell what is included in a specific function and/or to which categories a certain resource belongs. The next two subsections provide the function and resources classifications with definitions.

Assessment Function. Assessment functions are the discrete activities associated with an assessment strategy. Functions represent what an instructor, supervisor, or credentialing agent actually does in the course of assessing the student. Six assessment functions have been identified:

(1) Development or Revision (of instruments and procedures)
(2) Administration
(3) Analysis
(4) Management
(5) Advisement on Non-Competence
(6) Aggregate Judgment of Competence

These are defined as follows:

(1) Development/Revision. After the initial development of instruments and procedures, revision refers to the adjustments made of an assessment strategy prior to its use. For example, the questions on a test may need revision. The instructor may improve questions, test procedures, or format. Revision also includes the incorporation of new information. In fields where high technology or major philosophical shifts require frequent, drastic revisions, this function becomes a central cost.

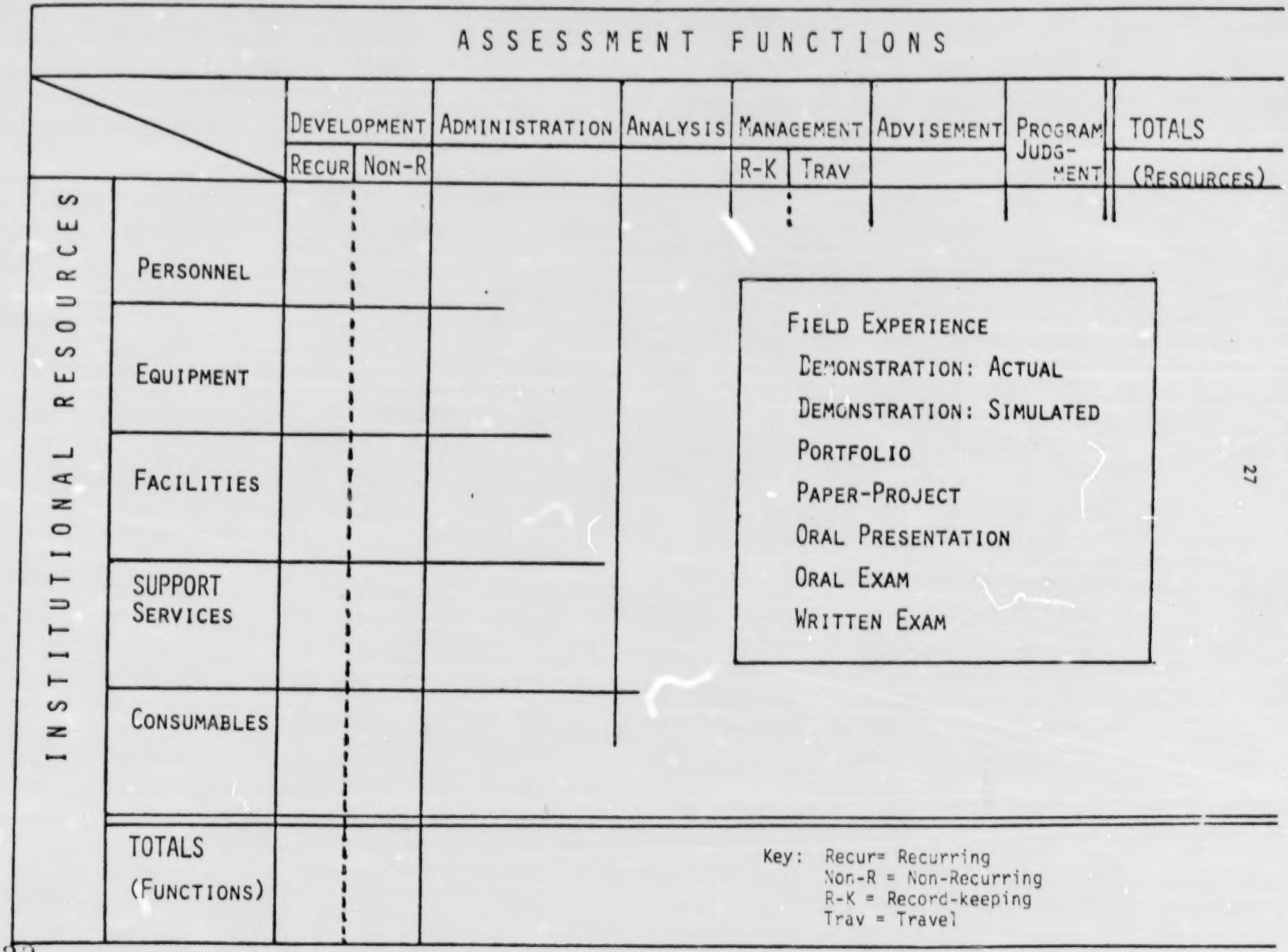



Fig. 5 - Primary Cost Analysis Matrix.

(2) Administration. Administration includes the time and procedures necessary to assess students. Traditionally, the standards for this function are the classroom time and activities used to give paper and pencil tests. Administration becomes more complex and costly as higher order assessment strategies are employed. Included in Administration are travel costs for site visits.

(3) Analysis. Analysis refers to the activities associated with examining, processing, and interpreting performance data from a particular strategy. Judgments of competence or non-competence are based on analysis. The most familiar form of assessment analysis is test correction and scoring. Less familiar forms of analysis are (a) observations related to field assessment, and (b) analysis of paper/projects, portfolios, and oral presentations.

(4) Management. The management function refers to the activities of recording and reporting data and analyses. Normally these activities relate to student grades. However, where instructors approach assessment data with an eye to developing or improving student competence, recording and reporting assessment data become crucial to providing valid diagnosis and prescription. Management also includes planning assessment programs for a semester.

(5) Advisement on Non-Competence. Advisement on non-competence involves all the activities of dealing with students who fail to achieve necessary levels of performance. CBE programs normally recycle these students, but only after advising them of error and misconception. In graduate level courses this function may lead to counselling students out of a program or field and, if appropriate, into another.

(6) Aggregate Judgment of Competence. This function refers to the summative judgment given on a student's cumulative performance at the conclusion of a program. It is the point at which the department, institution, or credentialing agency certifies the overall student competence. The function, as its name suggests, is the aggregation of all the separate assessments into some single, final decision that a student is or is not ready for professional practice.

Resources. As proponents of program budgeting know, identifying all the resources used solely for a particular program, (or, in this case, an assessment strategy) is not an easy matter. Institutional budgeting and accounting systems are generally based on jurisdiction rather than function/resource. Jurisdictional accounting, as the name implies, accounts for

the allocation and expenditure of funds by each organizational unit rather than for the actual function for which the funds are used. Jurisdictional, or line-item, accounting is too general and undifferentiated to use in attempting to determine how much a given program or set of assessment procedures will cost.

In order to identify the resources which go into a specific program, it is necessary to identify the various functions and then to specify what costs are constituted by resources used in these functions. Costing generally occurs, then, by estimating the percent of time for faculty and staff, the percent of use for facilities and equipment (where not totally allocated to a single program), and the cost of consumables. (Typically for education, the major costs are for faculty time.) These functional estimates are then translated into dollar costs, where possible, from accounting or budget figures.

Resources, then, are the entire variety of individuals and materials (people, supplies, equipment, and facilities) necessary and available for carrying out the assessment functions. Five resource categories are identified for cost analysis of assessment strategies.

(1) Personnel
(2) Equipment
(3) Facilities
(4) Support Services
(5) Consumables

They are identified as follows:

(1) Personnel. Personnel are those individuals who function within a discreet instructional unit such as an academic department, or in association with such a unit. Typically they are (1) instructors and field supervisors, (2) managers, and (3) support staff. The critical group consists of the instructors and field supervisors who handle

the actual student assessment. Therefore, an analysis of how instructors and field supervisors spend their time on student assessment is critical to a cost analysis of that activity. However, the contribution made to the assessment process by unit managers (e.g., chairpersons or deans) and secretarial and clerical staff must not be overlooked. Thorough analysis will consider the total contribution of all personnel with respect to the assessment process.

(2) Equipment. Equipment includes the durable, reusable hardware and software used in the assessment process. These may be books, computers, games, media equipment, etc. Equipment is distinguished from consumables because the former is used more than once.

(3) Facilities. Facilities are classroom, office, and other specialized areas necessary for assessment. When a facility is basic to all assessment strategies, it is not usually included in cost analysis. However, where a specific facility is necessary to only a certain type or class of strategies, it must be costed.

(4) Support Services. Support services refer to those services provided to aid the assessment process. Examples are (1) teaching assistants whose duties consist of administering and analyzing student assessment data, and (2) computer services used to analyze and report assessment data. Important new support services in assessment programs are those provided by assessment centers.

(5) Consumables. Consumables are those materials which are used during the assessment activity but which are not re-usable at its conclusion. Typically these include answer sheets for paper/pencil tests, record keeping booklets, computer printouts used in assessment, observation forms, etc.

Primary cost analysis thus involves the direct costing of function times resources times strategies and the eventual derivation of an actual dollar amount for each assessment strategy or combination of strategies in a given assessment system. A cursory inspection of the cost analysis matrix (Figure 5) may make the costing procedure appear deceptively easy. However, there are many problems, major and minor, that should be acknowledged in conducting cost analyses. Three cost-related concerns fundamental to the proper application of the model described here are:

(1) Distinguishing different types of costs
(2) Choosing the basis for cost comparison
(3) Accounting for variance in time and usage estimates

A brief description of each of these follows.

Types of costs. It is essential to be able to distinguish among different cost categories as confusion of one type of cost with another will distort the analysis and result in either under- or over-estimation of costs. Some of the basic costs which a cost analyst must be able to differentiate are fixed and variable costs, recurring and non-recurring costs, joint costs, and sunk costs.

Fixed costs are those costs which remain stable despite fluctuations in a program's operations. Most major capital expenses are regarded as fixed costs. Costs for facilities are usually fixed, since building mortgages and upkeep continue at a fixed level whether there are many or few students. Variable costs are costs which are directly related to the level of activity of the system. Thus, expenditures for materials are variable costs, since the more students there are, the more materials are required.

Recurring costs are those which occur during each and every cycle of a program's operation. Non-recurring costs are one-time costs during a program's life; for instance, costs for research and development. If these two types of costs are confused, especially when a non-recurring cost is listed as a recurring one, the entire program operation cost is inflated. In the cost matrix presented for use here, there are only recurring costs, since the analysis deals solely with operational costs of assessment systems.

Joint costs are those costs which must be allocated to more than one program or jurisdiction. They are costs which are shared. For example, different instructional programs may share the same equipment, or teachers may teach in more than one program. Determining percentages of use and time for such joint costs is difficult but crucial.

Sunk costs are those expenditures which have already been made at some time in the past. A common error in cost analysis and in decision-making is to include sunk costs in future projections of the cost of a program. Once resources have been irretrievably allocated, they should be disregarded in calculating program costs. This reflects the approach of using the "present worth" concept of economic analysis, which allows all costs and decisions to be based on the value or resources at the moment of analysis. The general principle behind ignoring sunk costs is to avoid committing resources which are still flexible in the present, on the basis of the past worth of resources which have been allocated and therefore, in effect, no longer exist. This procedure allows each decision to be based solely on the current value of flexible, available resources.

Basis of Cost Comparison. There are essentially two kinds of comparisons which a decision-maker may want to investigate. First, decision-makers may want to compare the absolute costs of various strategies or clusters of strategies in order to identify the general array of costs involved in every strategy or combination of interest. The focus is on general strategy comparison for which strategy comparable costs must be gathered. These costs include virtually everything which is necessary to any of the assessment strategies. There is one exception.

Where all strategies require the same basic resource (e.g., classroom space), then that resource need not be included since adding this cost merely inflates the cost of each strategy by a similar amount and gives an unduly high estimate for all strategies. Resources which can generally be eliminated from costing in this manner are basic resources which are fundamental to the enterprise of education and not unique to any particular program or activity.

The second type of comparison that decision-makers may want to investigate is one which indicates which strategy or group of strategies is most appropriate for a particular situation or setting. Program specific costs are appropriate for such comparisons. These costs are costs only for resources which are not already available to a program. Resources which a program already has would not be included in costing of a particular strategy. For instance, if a simulation strategy required use of an assessment center, and a center already existed and was accessible, then construction, maintenance, and staffing would not need to be included. Only the costs per use of the center would be figured into the strategy. However, if no such center existed, then all the costs associated with creating one would have to be included in the strategy cost analysis.

These two different comparison contexts--strategy comparable costs and program specific costs--are both important depending on the level of decison-making involved. If a federal or state agency is estimating average costs for a series of assessment systems, the former method may be more revealing in providing some sense of overall costs. If a specific program is attempting to determine its own system, the latter analysis is, no doubt, of more use.

Time and Usage Estimation. A major problem is the lack of validity and reliability data for the various estimates of time and usage which underlie the cost in the function/resource matrix for each strategy. However, there are ways to increase the validity and reliability of the estimates. First, clear definitions of functions will help the estimator (and the faculty member) distinguish activities (functions) and so estimate more accurately the proportion of time spent on each one. Second, redundancy of estimation in the form of multiple measures repeated over time will help identify any inconsistencies and provide some sense of the variance in responses. Third, well-conceived and validated measuring instruments and techniques will obviously increase confidence in the final estimates.

Rating scales, interviews, self-report plans, and institutional requirements are some of the typical procedures for gaining such information. More sophisticated techniques, for example, some of the procedures used in perceptual psychology, decision analysis, values clarification, and research, can be adapted when the costs are significant enough to warrant such an investment.

In the last analysis, a sensitivity analysis may reveal that over- or under-estimates of time may vary widely before any substantial effect on the choice of strategy occurs. For example, it is quite possible that any systematic over- or under-estimation of time or usage may occur across all strategies and thus affect all estimates equally. In sum, where the analysis seems grossly distorted, then experience, common sense, and intuition should be used to help give a more accurate reflection of true costs.

Secondary Cost Analysis. By and large, current cost analysis disregards non-material costs because it is difficult to assign numerical values to these costs. Secondary cost analysis, however, consciously considers these non-material costs. Typically these are social costs, psychological costs, lost opportunity costs, and the like. Because dollar values are not easily assigned to such costs, they tend to be overlooked and thus important decision-making information is discounted.

In fact, the designation "secondary cost analysis" is a bit misleading. The consideration of non-quantifiable costs and unintended outcomes is secondary only in the sense that the techniques for evaluating such costs and events are much less exact and less easy to apply. The concerns in secondary analysis are often more significant and have a greater impact on decisions than those dollar concerns considered in the primary cost analysis. This is particularly true when secondary analysis involves political and social issues which have ramifications for the public interest beyond the scope of dollars spent. For example, if a new competency-based program with high standards produces fewer graduates, albeit with better skills, there will be ramifications. There is a cost to the segment of the public which might have received some service, though of lesser quality, but must now do without any service while waiting for better trained graduates. On the other hand, there are a different set of costs when more graduates are produced but some are only capable of substandard service.

Some of the more fundamental secondary (or qualitative) costs for CBE are:

(1) Social costs
(2) Opportunity costs

(3) Accuracy and information costs
(4) Inflexibility costs
(5) Political costs

These will be discussed in turn below.

Social costs are those negative external outcomes which affect large segments of the public. Pollution, traffic congestion, and nuclear contamination exemplify social costs. Economists often refer to these as "externalities," a term used to designate costs of production not directly borne by the producer. Externalities may produce either positive or negative effects.

Opportunity costs are the cost of the next best alternative for which resources could have been used. This definition coincides with the economic definition of "cost" referred to earlier. It simply means that if the alternatives are to select a performance-based CBE assessment system or a new information management program, the cost of the assessment system is conceived as the cost of the new program, the next best alternative given up or foregone. Expressing costs as opportunity losses is often much more dramatic than simply presenting a dollar expenditure amount because it forces people to confront graphically the things they are giving up when they choose one alternative over another.

Accuracy and information costs refer to the cost of acquiring additional information in order to be able to make a more accurate judgment. Increased accuracy requires increased information which entails increased cost.

Inflexibility costs are those costs associated with decisions which cannot be easily reversed. Often resources must be committed irretrievably in choosing one program over another or one approach to strategy

assessment over another and not much can be salvaged if a later decision is made to abandon the project.

Finally, political costs are those costs which occur generally to decision-makers, and which involve losses in power, status, position, or office. The most obvious case of such costs is one where a politician supports an unpopular law and is not reelected because of it. This is not a trivial factor inasmuch as much of the drive behind CBE is political rather than (or perhaps as well as) educational. Thus the political costs of adopting or not adopting competency programs and assessments should not be ignored in the activity of cost analysis.

Each of these types of cost resists easy quantitative analysis which would permit an assignment of a dollar value to them. Therefore, they are often overlooked or sidestepped. One tactic in cost-effectiveness studies is to view such costs as "negative" effectiveness and to let the effectiveness analysts wrestle with the difficulties of measurement. Our position is that the secondary costs are too important to be neglected and thus hospitality to these qualitative factors has been provided in our model.

Even though dealing with these secondary costs is difficult, we can suggest some guidelines. First, when dealing with costs which are not easily quantified--for instance, the costs to students of faculty evaluation which mistakenly eliminates some competent individuals from a program versus the costs to society of faculty evaluation which passes non-competent students into public practice--the various costs can at least be ranked according to their perceived importance. Thus, given limited resources, if it is felt that society's well-being has relatively more value than that of each student, then a system can be designed which

requires more expert performance and is less lenient in borderline cases even though no dollar cost is placed on such values. (In practice, dollar amounts can often be found for such costs through the measurement of expected utilities of decision-makers or of certain segments of society. See the decision analysis literature for a more complete account of these techniques.)[53]

A second technique for dealing with secondary costs is to get some estimates of how important these factors are in the overall decisions to be made. In the above example, one approach might be to determine how much a possible student or public lawsuit might cost an institution if it were determined that the assessment strategies for determining competence were either too stringent or too lenient, and what the possibilities of such judgments might be. Given these estimates, an analysis could be made which would reveal whether different strategies or standards should be used, and at what point additional costs for better assessment cease to balance the costs to students and society. In other words, if a student lawsuit might cost an institution $10,000, and the likelihood of such a suit was 1 in 500, then the expected cost would be $10,000 x .002 - $20 per 500 students trained. If the cost of assessment to reduce the likelihood of such an occurrence to 1 in 1000, or .001, was $20 per student, or $10,000 for 500 students, then investment would not be reasonable.

Many people are extremely uncomfortable putting questions of value into monetary terms, however vague. Still it is essential to realize that putting dollar costs on values such as "fairness to students" or "responsibility to society" occurs at every junction of the academic process where allocations of scarce funds are made to one type of activity and not another. The argument above is not meant to indicate that

student concerns are unimportant or always of lowest priority. Instead, it emphasizes rather dramatically that tradeoffs are fundamental to most decisions, and it forces decision-makers to understand that the abstract idea of assessing competence has very real and often contradictory consequences for students and the public.

Although these secondary costs should not be ignored, it is generally not worth the investment to attempt to be too refined in placing monetary values on all costs or anticipating fully all outcomes. More often than not, an awareness of these costs and potential effects may be sufficient.

The two primary dimensions of the model--cost and effectiveness--having been discussed in some detail, the components can now be brought together in a discussion of the model for cost-effectiveness analysis of assessment strategies for competency-based education.

## The Cost-Effectiveness Model

Figure 6 provides a graphic overview of the model. It is drawn to illustrate the way in which the cost-effectiveness analysis of assessment programs is to be approached. It is once a cost analysis and again an effectiveness analysis. The two analyses are undertaken for the purpose of exposing differences between alternative assessment strategies. The data developed are then recapitulated in summary form in such a fashion that the decision-maker can focus on the pertinent and relevant facts in order to make informed decisions and reasonable policies.

The bulk of the analysis effort is concentrated in the center of the model and is represented by the four blocks:

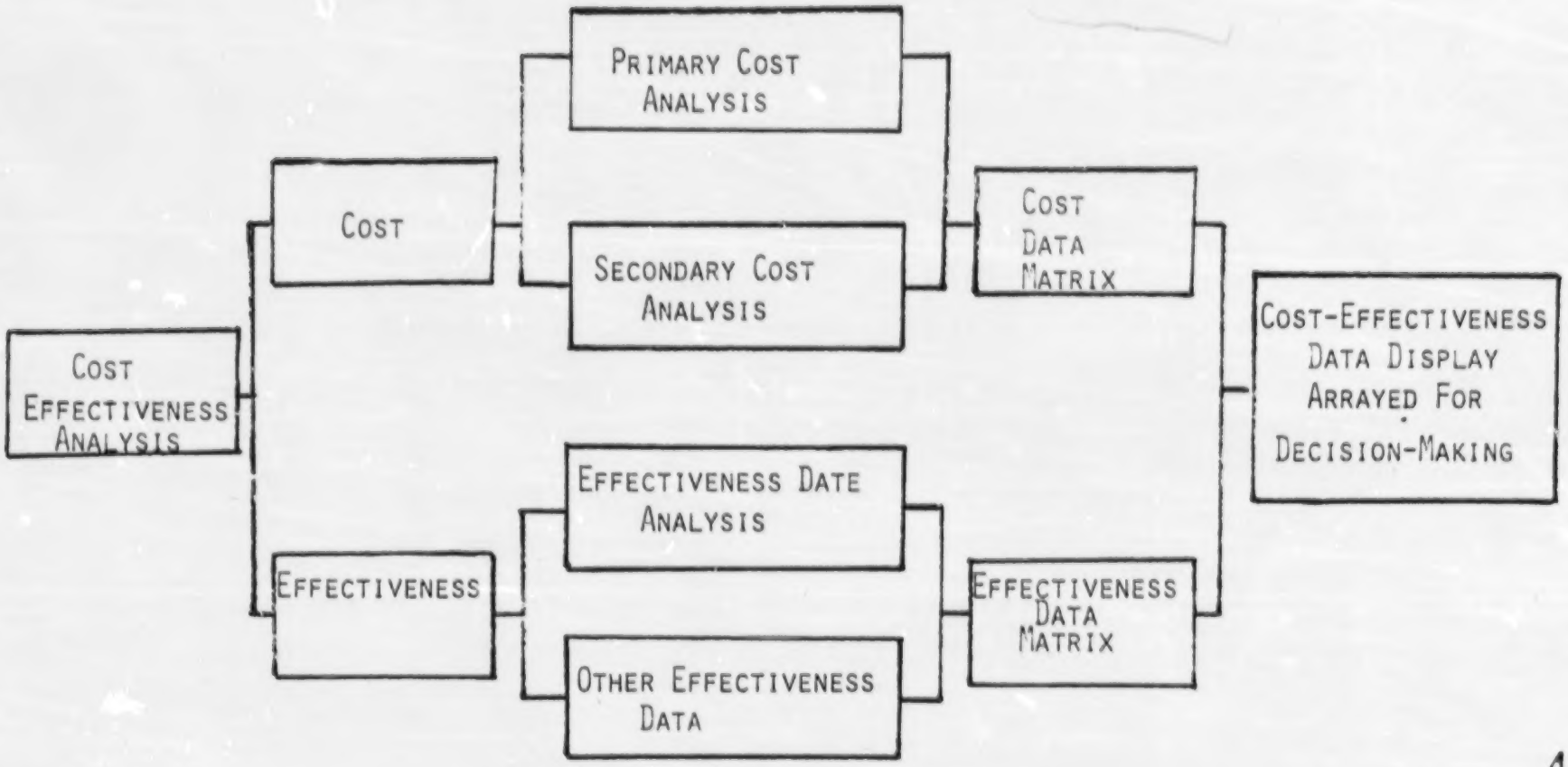



Fig. 6. General Cost-Effectiveness Analysis Model for CBE Assessment Programs.

(1) Primary cost analysis
(2) Secondary cost analysis
(3) Primary effectiveness analysis
(4) Secondary effectiveness analysis

The results of the first pair of analyses are placed in the Cost Data Matrix and those of the second pair in the Effectiveness Data Matrix. The data from these matrices are condensed again and brought together in the Data Display for Decision-Making.

The dichotomous approach--cost and then effectiveness--is used for the sake of completeness. It has the advantage of pointing up to analysts what they typically overlook in approaching cost-effectiveness analysis. The division into primary and secondary parts, more or less corresponding to a division into quantitative and qualitative measures, gives the decision-maker an opportunity to see what is critical and what is less so. The final choice then should be based on a holistic view of the situation, even though circumstances may dictate against detailed secondary forms of analysis. Here, however, the decision-maker can control the analysis process by selecting which elements of secondary analysis will stay in the final array and which will be omitted. The detailed description of the elements suggested for inclusion in the secondary analyses (given in the two previous sections above) should assist in making analysts and decision-makers fully aware of the reasons why certain secondary analysis has not occurred.

The Matrices. Cost analysis and effectiveness analysis are greatly facilitated by the use of matrices. Doughty and Stakenas make a strong recommendation for their use in reporting and relating information for decision-makers.[54] In addition to their value in communicating informa-

tion, matrices are useful as a format in which to arrange data en route through the analysis. Often matrices become part and parcel of the analysis process and may even serve to guide the analysis itself. Matrices are used in the present model for all these reasons.

The Cost Data Matrix is used for three reasons: (1) to guide cost analysis, (2) to display and relate information for decision-makers, and (3) as a residual for the placement of available data. For the present the Cost Data Matrix is confined largely to primary cost analysis. The growth and development of the cost analysis of competency-based assessment programs may expand the use of this matrix to secondary cost analysis as well. The reader is referred to Figure 5 for details on the Cost Data Matrix.

The Effectiveness Data Matrix is used in two ways: (1) as a residual for effectiveness data, and (2) as a reporting device. Principally the Effectiveness Data Matrix serves to display effectiveness data for analysis and interpretation. Unlike the cost matrix, however, the effectiveness matrix is not confined to primary analysis. Still, the degree to which it will include secondary forms of effectiveness analysis data is highly situational. Because an effectiveness analysis is essentially a validity analysis, the data from such an analysis may not require a matrix. The Effectiveness Data Matrix is recommended, nevertheless, because its use indicates to the evaluator the breadth of the analysis and it forces the evaluator to ask secondary analysis questions. Figure 3 provides the detail on the Effectiveness Data Matrix.

Data Display for Decision-Making. The final major element in the general model is the Data Display for Decision-Making. For the display

of analysis data, Doughty and Stakenas suggest a tabular array approach.[55] Their claim is that such an approach

> enables comparisons to be made of different kinds of quantitative data as well as qualitative criteria. Single dollar cost measures of an alternative are as uncommunicative and uninformative as any contrived single measure of effectiveness.[56]

They emphasize the need to let the array "reflect a set of cost and effectiveness measures suitable for use in making decisions."[57] They further suggest a two-dimensional field for comparison, which in this case would be Strategy by comparative Cost Effectiveness criteria, or

Assessment Strategy x Cost-Effectiveness Data

Figure 7 illustrates this arrangement.

Fixing Costs or Fixing Effectiveness. In using the model it is probably necessary to "fix" at least one of the two standards of comparison: cost or effectiveness. Fixing costs or effectiveness means setting some minimally acceptable level of effectiveness, or some maximally acceptable level of cost, and allowing the unfixed factor to vary. In comparing systems of assessment strategies, if there is a fixed resource limit beyond which costs cannot range, then the effectiveness of systems whose costs are under this ceiling is the criterion on which a selection will be made. Conversely, if a certain level of achievement is essential, for instance, reading at the national norm for grade level, then cost becomes the principal factor in choosing among programs which reach this achievement level or above.

At this point the cost-effectiveness model and its components have

| Strategies / Criteria | Field Experience | Actual Demonstration | Simulated Demonstration | Portfolio | Papers-projects | Oral Presentations | Oral Exams | Written Exams |
|---|---|---|---|---|---|---|---|---|
| Effectiveness: | | | | | | | | |
| Primary | | | | | | | | |
| Secondary | | | | | | | | |
| Cost: | | | | | | | | |
| Primary | | | | | | | | |
| Secondary | | | | | | | | |







Fig. 7. Data Display for Decision-Making

been described. However, it is one thing to develop an ideal, theoretical model in the abstract and quite another to apply that model in a real situation. In the next section, some preliminary attempts to apply one aspect of the model through collecting primary cost data from faculty will be described.

General Competency Assessment by Faculty

There were two specific purposes in this portion of the Competency Assessment study. First, examination of how effectively competencies were presently being assessed within the regular courses in a library school and in a school of education needed to be determined. Secondly, some baseline cost data needed to be collected both to test the primary cost matrix of our model and also to begin the process of building an empirical data bank of cost information. This section of the study was guided by four research questions:

(1) What is the distribution of competencies within courses?

(2) What is the distribution of assessment strategies within courses?

(3) What is the cost of course-related assessment activities?

(4) How is the effectiveness of assessment strategies to be determined?

Each of these questions will be discussed in sequence.

Distribution of Competencies within Courses. A year and a half earlier all the faculty of the School of Information Studies and of the Area of Instructional Technology, School of Education were interviewed

to determine which competencies were taught in their classes in order to advise students which particular set of courses to take. Advisement followed a thorough diagnosis of each student's present level of skills and abilities with a joint determination of where the student needed to become more proficient. After the determination of a desired set of competencies to be acquired or added to, the courses were then examined to see which group of courses would be the most appropriate for the student.

During this part of the project, the researchers went back to the faculty to ask them not just which competencies they taught, but which ones were actually assessed and in what manner the assessment took place. Outlines, activities, and assignments were collected from faculty for each relevant course and then classified by types of assessment procedures. The interviews were analyzed and a new competency-by-course matrix drawn (see Figure 8). Not surprisingly, it was found that the faculty teach more competencies than they actually assess.

A basic assumption that guided the work throughout the project is that the school media competency program must fit within a larger library school context composed of courses, many of which are general in nature, and directed to students who will work in a variety of library settings. The importance of this approach should be underscored. The alternative assumption would be that the school media program would be treated as a closed system isolated and apart. Under this assumption, there would be a school within a school with a set of courses tailored to fit the school competency-based program, and which prospective school media specialists would take by themselves away from the rest of the school.

The broader approach has a number of advantages. It allows school media specialist students to relate their work to the larger context.

| Competency Number | ITE 501 | ITE 511 | ... | ITE 719 | IST 502 | IST 503 | ... | IST 657 |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | X | | | | |
| 2 | | X | | | | | | |
| 3 | | | | | | X | | |
| 4 | | | | | | X | | |
| 38 | X | | | | | | | |
| 39 | X | | | X | | | | |
| 40 | | | | | | | | |
| 69 | | | | | | | | X |
| 70 | | | | X | | | | X |

Fig. 8. Course by Competency Matrix.

Of equal or greater importance, it enlarges faculty awareness of the special needs and contributions being made by the modern school library media center. This positive benefit more than offsets the requirement for regularly checking with all faculty--old and new--for changes in courses, new content, new methods, etc. Competency assessment then becomes a dynamic process continually being adjusted to the changing issues and the changing environment.

Distribution of Assessment Strategies within Courses. In the re-interviews with faculty it was found that there were redundancies--many competencies taught and assessed from different points of view in different courses. There were also some gaps where competencies were not assessed at all through formal courses. Figure 9 shows the ranking of frequency of assessment by competency number. Each year this analysis needs to be a part of the process. Those competencies not being directly assessed through course work may be assessed through field work, independent projects, or on an individual or small group basis outside of courses. As a result of the analysis, suggestions will be made for course revision to include and/or to measure the attainment of other competencies where appropriate.

Data did not always fall out as neatly as the researchers might have liked. In some cases the data revealed that the courses allowed students to demonstrate only a few of the competencies, while in other cases many competencies could be demonstrated. In all cases, however, the courses considered more than the competencies of interest to the project and the faculty were assessing more than just those competencies. This is important as it means the competency-based program sets a floor

| Statement Number | |
|---|---|
| 22 | - 7 |
| 15<br>16<br>19<br>20<br>29 | - 6 |
| 1 | |

| Statement Number | |
|---|---|
| 32<br>33<br>42<br>45<br>46<br>47 | - 2 |
| 1 | |
| 37<br>39<br>50<br>52<br>53<br>62<br>66<br>67<br>68 | - 0 |

49

Fig. 9. Frequency of Competency Assessment through Courses.

rather than a ceiling. Many qualifying statements accompanied faculty responses. These statements explained the use of a given assessment technique. They emphasized the degree to which the technique assessed a given competency. And they revealed the instructor's approach to assessment, which led to an exploration of the impact of differing faculty styles of assessment.

Cost and Effectiveness of Course-Related Assessment Activities. It is necessary to collect specific cost and effectiveness data on a periodic basis and to compare results in order to make responsible resource allocations satisfying through sub-optimization the needs of the program, the school, and society. The researchers have carefully developed a model that specifies the universe of cost categories and effectiveness criteria.

Three alternatives were considered for testing the cost model. It is possible to look at costs of field work; secondly, an assessment situation could be established outside of classes through simulation, laboratories and summative testing; or an attempt could be made to cost the assessment activity that was already going on in classes. The third alternative was the most appealing since it probably would have the broadest applicability to other programs.

Only primary cost information was collected and no attempts were made to grapple with social costs, psychological costs, lost opportunity costs, and the like. Nor would data be collected to determine degree of effectiveness at this time. A simplifying assumption was adopted, i.e., that assessment through course work was primarily a function of faculty time. Faculty time was assumed to be differentially allocated to assess-

ment activity based on: (1) whether or not the course had been taught previously, (2) the number of students in the course, (3) the number of competencies being assessed in the course, and (4) the faculty style of assessment (including the number of assessment activities).

Using these assumptions, the researchers asked faculty members to estimate the number of hours spent on assessment activity for a given course, the proportion of that assessment time dedicated to the specified competencies relevant to the school media program, and the estimated accuracy of the estimates. Figure 10 shows the data collected.

It was decided not to try to collect data retrospectively in the various function categories of test administration, test analysis, test recording-reporting, and test-related advisement, but rather to be satisfied with a global measure, as it was felt that the accuracy level of the estimates was going to be somewhat low and it would further attenuate the accuracy to break estimates into more discrete categories.

The faculty were also asked to estimate the costs of material used in the assessment activity and any other costs associated with assessment, for example, travel, graduate assistant time, equipment, etc. The time estimates were converted to dollars using a conservative factor (based on average annual professorial salaries divided by an estimated fifty hour work week) of $15 per hour. Cost of materials and other costs were also converted to dollar amounts.

Figure 11 shows a sample of the results collected in this pilot test. The two courses shown were very different in nature. ITE 719 is an administration course; IST 612 is a children's literature course. Although more competencies were assessed in the administration course, there were twice as many students in the literature course and nearly three times

1. Course number

2. Number of Competencies Available in Course

3. Number of Functional Areas Involved in Course

4. Assessment Strategies Employed in Course

5. Ratio of Number of Competencies Assessed To Number of Competencies Available

6. Actual Number of Hours Spent in Assessment Activity by Faculty Member

7. Estimate of Costs Other than Faculty Time

8. Level of Accuracy of Estimates

| 0 | 1 | 2 |
|---|---|---|
| Wild Guess | Ballpark Figure | Accurate Estimate |

Fig. 10. Cost Data Collected from Faculty.

| COURSE | HOURS | COST OF TIME | OTHER COSTS | TOTAL COST |
|---|---|---|---|---|
| ITE 719 | 31 | $465 | $30 | $ 495 |
| IST 612 | 84 | $1260 | $30 | $1290 |

| | | | |
|---|---|---|---|
| RANGE OF COSTS | $ 225 | TO | $ 1470 |
| MEAN | $ 804 | | |
| MEDIAN | $ 915 | | |
| TOTAL COSTS | $ 19,283 | FOR 24 | COURSES |

Fig. 11. Cost of Assessment-Related Activities: Exploratory Study

the number of assessment activities undertaken. These factors account for most of the difference in total cost. The range of cost of the twenty-four courses for which data was collected varied widely although the mean and median measures were fairly close.

Obviously, only limited conclusions can be reached from this small retrospective data gathering effort. Many more factors would need to be included and real-time measures applied. The exercise did, however, indicate the feasibility of the approach and the willingness of the total faculty to assist in the data collection effort.

## Elaboration of the Cost-Effectiveness Model

Several specific areas are in need of elaboration:

(1) The specific procedures for collecting and calculating cost data need greater specification. Especially needed are techniques for estimating average time and usage estimates, translating traditional budgets to function and resource categories, and deriving program specific costs.

(2) The relationship of the effectiveness model to test validation procedures needs to be thoroughly explored. Validation procedures for non-pencil-and-paper tests need to be identified and/or developed.

(3) The implied correlations between strategies and degree of effectiveness in assessing competence must be experimentally validated.

## REFERENCES

[1]Daniel, Evelyn H. and Donald P. Ely. A Process for Developing a Competency Based Educational Program for Media Professionals. (Syracuse: ERIC Clearinghouse on Information Resources, 1977. ED 149 740)

[2]Daniel, Evelyn H. and Donald P. Ely. Developing Alternative Strategies for Evaluating Competencies of Library/Media Personnel in Competency Based Programs. A Research Proposal to U.S. Office of Education. (Syracuse: Syracuse University, 1978).

[3]Cyrs, T. E. and D. J. Dobbert. A Competency-Based Curriculum: What is it? (Minneapolis: University of Minnesota, College of Pharmacy, 1976).

[4]Chisholm, Margaret and Donald P. Ely. Media Personnel in Education: A Competency Approach. (Englewood Cliffs, N.J.: Prentice-Hall, 1976).

[5]American Association of School Librarians. Committee on the Certification of School Media Specialists. Certification Model for Professional School Media Personnel. (Chicago: American Library Association, 1976).

[6]Schalock, D., "Cost and Benefits," pp. 114-166 of From Commitment to Practice: The Oregon College of Education, Elementary Teacher Education Program. (Salem, Ore.: American Association of State Teacher Education Programs, 1978).

[7]Okey, J. R., W. Capie, and C. E. Johnson. Teacher Performance Assessment Instruments: Manual for Data Collectors. (Atlanta, Ga.: Georgia State Department of Education, 1978).

[8]Ellett, D. D. and C. E. Johnson. A Study of the Construct Validity of the Teacher Performance Assessment Instruments: Comparable Instruments Investigation. A Working Paper. (Athens, Ga.: Georgia University, College of Education, July 1978).

[9]Williams, R. G. Competency Maintenance System Project. Final Report. (Carbondale, Ill.: Southern Illinois University, School of Medicine, 1976).

[10]University of Minnesota, College of Pharmacy. Competency Identification Including Evaluation and Validation. Working Paper no. 2. (Minneapolis: University of Minnesota, College of Pharmacy, Curriculum Improvement Office, June 1, 1976).

[11]University of Washington. Consortium-Directed Competency-Based Program. (Seattle: University of Washington, 1976?).

[12]Ellis, J. A. and J. E. Bryant, "Competency-Based Certification for School Social Workers," Social Work 21 (1976) 381-385.

[13]"College Degrees Awarded on Basis of Competency-Based Education," Competency Forum 2 (March 1978) 12-14.

[14]Reported at program on "Performance/Criterion Evaluation: How to Tell the Competent from the Non-Competent" at the April 1979 Association for Educational Communications and Technology Conference in Kansas City.

[15]Scanland, Worth, "Criterion Referenced Measures and Performance Based Measures." (Pensacola: Naval Education and Training Command, 1979). Reported at 1979 AECT Conference.

[16]Waag, Wayne, "Development of an Objective Performance Measurement System for Aircrew Assessment." (Williams Air Force Base, Ariz.: Air Force Human Resources Laboratory, 1979). Reported at the AECT Conference.

[17]Wilshire, Robert, "Description of the SQT (U.S. Army Skill Qualification Test); Breakthrough in Performance Testing." (Ft. Gordon, Ga., U.S. Army Signal Center, 1979). Presented at 1979 AECT Conference.

[18]Kupper, Marjorie, "The Design of a Hands-On Performance Criterion Test." (Fort Belvoir, Va.; U.S. Army Engineer School, 1979). Reported at the AECT Conference.

[19]DeProspo, Ernest R. and James W. Liesener, "Media Program Evaluation: A Working Framework," School Media Quarterly 4 (Summer 1975) 288-318, 323-336.

[20]Hall, Gene E. and Howard L. Jones. Competency-Based Education: A Process for the Improvement of Education. (Englewood Cliffs, N.J.: Prentice-Hall, 1976).

[21]Houston, W. R. and R. B. Howsam, eds. Competency-Based Teacher Education: Progress, Problems, and Prospects. (Chicago: Science Research Associates, 1972).

[22]Hodgkinson, H., J. Hurst and H. Levine. Improving and Assessing Performance: Evaluation in Higher Education. (Berkeley, Ca.: Center for Research and Development in Higher Education, 1975).

[23]Harris, J. and S. P. Kelly, "Assessment Measures Needed for Competency-Based Higher Education," Peabody Journal of Education 53 (1977).

[24]Grant, Gerald. On Competence: A Critical Analysis of Competency-Based Reforms in Higher Education. (San Francisco: Jossey-Bass, 1979).

[25]King, Susan E., "Assessment of Competence: Technical Problems and Publications," in On Competence, ed. by Gerald Grant, pp. 491-521. (San Francisco: Jossey-Bass, 1979).

[26]Gilbert, Thomas F. Human Competence: Engineering Worthy Performance. (New York: McGraw-Hill, 1978).

[27]Stufflebeam, Daniel L., "Evaluation as Enlightenment for Decision-Making," in Improving Educational Assessment and an Inventory of Affective Behavior, ed. by W. H. Beatty. (Washington, D.C.: Association for Supervision and Curriculum Development, National Education Association, 1969).

[28]Rountree, Derek. Assessing Students: How Shall We Know Them? (New York: Harper & Row, 1977).

[29]Kay, R. M., "Measurement Techniques: What We Have and What We Need," in Defining Teaching Competency: Emerging Issues in the Development Period of the Assessment Approach to Competency-Based Teacher Education. (Washington, D.C.: Association for Educational Research, 1973).

[30]Moses, Joseph L. and William C. Byham, eds. Applying the Assessment Center Method. (New York: Pergamon Press, 1977).

[31]Hodkinson, H., J. Hurst and H. Levine. Ibid., p. 3.

[32]Rogers, Joanne V., "Media Competence of Teachers: A Review of Measurement Literature," Educational Technology 18 (October 1978) 16-22.

[33]Jobs in Instructional Media, ed. by C. James Wallington, et al. (Washington, D.C.: Association for Educational Communications and Technology, 1970).

[34]Mehrens, W. A. and I. J. Lehmann. Measurement and Evaluation in Education and Psychology. (New York: Holt, Rinehart and Winston, 1973).

[35]Holloway, Robert. University of Minnesota, School of Pharmacy. Personal Interview on their Competency-Based Education Program, October 20, 1978.

[36] Kendall, M. G. and W. Z. Buckland. A Dictionary of Statistical Terms. 2nd ed. (New York: Hafner, 1967).

[37] Mehrens, W. A. and I. J. Lehmann. Ibid., p. 128.

[38] Kendall, M. G. and W. Z. Buckland. Ibid., p. 309.

[39] Harris, J. and S. P. Kelly. Ibid.

[40] Doughty, Philip, et al. Guidelines for Cost-Effectiveness Analysis for Navy Training and Education. (San Diego: Navy Personnel Research and Development Center, June 1976).

[41] Knapp, J. and A. Sharon. A Comprehension of Assessment Techniques. (Princeton: Educational Testing Service, 1975).

[42] Kelly, E., "Evaluation Techniques: An Overview" Lecture. (Syracuse: Syracuse University, Fall 1977).

[43] Mehrens, W. A. and I. J. Lehmann. Ibid.

[44] Scriven, Michael. "If the Program Is Competency-Based, How Come the Evaluation Is Costing So Much?", in Competency Assessment Research and Evaluation: A Report of a National Conference. (Syracuse: The National Dissemination Center, Syracuse University, 1974).

[45] Scriven, Michael. Ibid., p. 161.

[46] Fisher, G. H. Cost Considerations in Systems Analysis. (New York: American Elsevier, 1971).

[47] The Resource Approach to the Analysis of Educational Project Cost. Number 3 in a Series of Monographs on Evaluation in Education. (Washington, D.C.: U.S. Dept. of Health, Education, and Welfare, 1978).

[48] Doughty, Philip and Al Beilby. Cost Analysis and Teacher Education: A Comment on Relevant Relationships and A Review of Existing Models. (Albany: State Education Department, Division of Teacher Education and Certification, 1974).

[49] Carpenter, Margaret B. and Sue A. Haggart, "Cost-Effectiveness Analysis for Educational Planning," Educational Technology (Oct. 1970) 26-30.

[50] Stace, P. Program Cost Evaluation: An Analysis of the Costs Incurred by Sponsoring a Program of Study Abroad. Unpublished doctoral dissertation. (Syracuse: Syracuse University, 1979).

[51] Carpenter, Margaret B. and Sue A. Haggart. Ibid.

[52] Doughty, Philip et al. Ibid.

[53] Brown, Rex V. Decision Analysis: An Overview. (New York: Holt, Rinehart, and Winston, 1974).

[54] Doughty, Philip and P. Stakenas, "An Analysis of Cost and Effectiveness of an Individualized Subject Offering," in Accountability: Systems Planning in Education, ed. by C. Sabine, pp. 175-194. (Homewood, Il.: ETC Publications, 1973).

[55] Ibid., p. 182.

[56] Ibid.

[57] Ibid.